Topology-Aware Physics-Informed Reinforcement Learning for Resilient Operation and Adaptive Restoration of Renewable-Integrated Active Distribution Networks

VanDu Nguyen

Faculty of Mechanical Engineering, Thu Duc College of Technology, Ho Chi Minh City, Vietnam.

DOI: https://doi.org/10.51583/IJLTEMAS.2026.150900003

Received: 17 September 2026; Accepted: 22 September 2026; Published: 29 September 2026

ABSTRACT

The increasing penetration of photovoltaic generation in active distribution networks (ADNs) introduces significant operational variability and challenges for resilient network restoration. This paper proposes a Topology-Aware Physics-Informed Reinforcement Learning (PIRL) framework for adaptive operation and self-healing of renewable-integrated ADNs. The proposed approach integrates topology-aware network representation, physics-informed reward guidance, coordinated distributed generation (DG) and battery energy storage system (BESS) control, and feeder reconfiguration within a unified learning framework. The controller is evaluated on the IEEE 33-bus ADN under different combinations of PV generation, load demand, and feeder contingencies. Simulation results show that PIRL achieves a minimum voltage of 0.976 p.u. and reduces network power loss to 121.3 kW. Under a feeder fault, the proposed framework restores the interrupted load within 13 min with an energy not supplied of 72 kWh and achieves a resilience index of 0.96. In addition, PIRL reaches stable policy convergence at approximately 560 training episodes with a final normalized cumulative reward of 0.91. These results demonstrate that combining physical operating information with adaptive reinforcement learning and coordinated DG-BESS control can improve voltage security, network efficiency, restoration capability, and learning efficiency in renewable-integrated ADNs.

Keywords: Physics-Informed Reinforcement Learning, Active Distribution Networks, Self-Healing, Battery Energy Storage Systems, Feeder Reconfiguration.

INTRODUCTION

The increasing deployment of renewable energy resources is transforming the operational characteristics of modern distribution networks. Distributed photovoltaic generation, inverter-based resources, and battery energy storage systems (BESSs) provide additional flexibility and create new opportunities for improving the efficiency of active distribution networks (ADNs). At the same time, the variability of renewable generation and load demand introduces continuously changing power injections and bidirectional power flows. These effects can complicate voltage regulation and increase the sensitivity of the network to operating disturbances (Byeon & Kim, 2024). The resulting uncertainty becomes more pronounced when the network must respond to rapidly changing renewable output and demand conditions (Huang et al., 2022), (Qiu et al., 2024).

Restoration following a feeder disturbance presents an additional layer of difficulty because the network topology may no longer remain unchanged. Once a feeder fault occurs, the affected section has to be isolated and alternative electrical paths may need to be established through switching operations. At the same time, distributed resources must be coordinated to support the restoration process while maintaining acceptable voltage and loading conditions. Existing approaches based on deterministic optimization and model predictive control can provide effective solutions when network states and operating conditions are sufficiently known (Glavic, 2019), (Byeon & Kim, 2024). Their application becomes more challenging, however, when renewable generation, load demand, and feeder connectivity vary simultaneously. Conventional restoration schemes may also rely on predefined switching procedures or centralized decision

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

structures, which can reduce their adaptability under changing contingency conditions (Glavic et al., 2021), (Zhou et al., 2021).

Deep reinforcement learning (DRL) has attracted increasing interest as an alternative approach for adaptive power-system control. Rather than relying exclusively on an explicitly formulated control sequence, a reinforcement learning agent can improve its policy through repeated interaction with the operating environment. This characteristic has motivated applications in energy management, voltage control, and distributed resource coordination (Z. Zhang et al., 2021), (Zhao et al., 2016). Several studies have demonstrated the potential of learning-based methods for controlling renewable-integrated distribution systems (B. Zhang et al., 2024), (Xu et al., 2017), (Zhao et al., 2016).

Despite these developments, two issues remain important. First, many learning-based formulations describe the network mainly through numerical state variables and do not explicitly exploit the structural relationships among buses and feeder branches. Such a representation may not adequately distinguish between operating states having similar electrical measurements but different network configurations. Second, purely data-driven exploration does not necessarily guarantee that the resulting actions satisfy the physical requirements of the distribution system. This can lead to undesirable voltage conditions, excessive losses, unnecessary switching, or degraded performance when the operating condition differs from the training environment (Li & Xu, 2024), (Wang et al., 2025).

The graph structure inherent in an ADN provides a natural means of addressing the first issue. Buses can be represented as graph nodes, while feeder branches form the corresponding edges. Electrical measurements can then be interpreted together with the network connectivity through which power flows between different parts of the system. Recent research has explored graph-based learning to capture interactions among network components, while physics-guided learning has been introduced to improve the physical consistency of intelligent control decisions (B. Zhang et al., 2024), (Wang et al., 2025). In distribution systems, such physical guidance is particularly important because voltage limits, feeder loading, power balance, and energy-storage operating constraints must remain satisfied during both normal operation and restoration (B. Zhang et al., 2024), (Dominguez-Garcia et al., 2024).

The combination of topology-aware learning and physical operating information is therefore particularly relevant to resilient ADN control. Following a disturbance, the controller must not only determine suitable control actions but also respond to the modified feeder connectivity. DG and BESS units can provide local voltage support, active-power flexibility, and restoration assistance, but their effectiveness depends on how they are coordinated with the prevailing network configuration. Existing resilience-oriented optimization and artificial-intelligence approaches have made considerable progress, although renewable uncertainty, distributed coordination, and rapid restoration remain challenging issues (Glavic et al., 2021), (Behzadi et al., 2024).

To address these issues, this paper develops a topology-aware Physics-Informed Reinforcement Learning framework, denoted as PIRL, for renewable-integrated active distribution networks. The framework combines graph-based network representation, reinforcement learning, physical operating constraints, coordinated DG-BESS control, and feeder reconfiguration. The graph component provides information about the electrical relationships among buses and feeder sections, whereas the physics-informed mechanism evaluates the consequences of learned actions according to the operating requirements of the ADN. The resulting controller is intended to adapt to renewable variability and changing feeder configurations while maintaining voltage security and restoration feasibility.

The main contributions of this study are as follows:

i) A topology-aware Physics-Informed Reinforcement Learning framework is developed for resilient operation of renewable-integrated ADNs.

ii) An electrical graph representation is incorporated into the learning process to retain the connectivity relationships among buses and feeder sections under changing network configurations.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

iii) A physics-guided action evaluation mechanism is introduced to connect learned control decisions with voltage, feeder loading, BESS operating limits, network losses, ENS, and switching requirements.

iv) DG reactive-power control, BESS operation, and feeder reconfiguration are coordinated within a common control framework to address both normal operation and post-fault restoration.

v) The proposed approach is evaluated using the IEEE 33-bus ADN under renewable variability, load changes, and feeder contingencies, with performance assessed in terms of voltage regulation, power loss, restoration capability, and learning convergence.

System Model and Problem Formulation

Distribution Network Model

The investigated system is based on the IEEE 33-bus active distribution network equipped with photovoltaic distributed generation and BESS units. The network represents a renewable-rich operating environment in which variable DG output and time-dependent demand can produce changing power-flow patterns and, in some operating conditions, bidirectional flows (Byeon & Kim, 2024), (Qiu et al., 2024). Consequently, voltage regulation, loss minimization, and reliable service restoration are considered important operating requirements.

The steady-state electrical behavior is described through the active- and reactive-power balance equations. The contributions from DG and BESS units must satisfy the network demand while remaining within the electrical limits of the distribution feeders.

$$P_i^{DG} + P_i^{BESS} - P_i^{load} = V_i \sum_{j=1}^{N} V_j\left(G_{ij}cos\theta_{ij} + B_{ij}sin\theta_{ij}\right) \tag{1}$$

$$Q_i^{DG} + Q_i^{BESS} - Q_i^{load} = V_i \sum_{j=1}^{N} V_j\left(G_{ij}cos\theta_{ij} + B_{ij}sin\theta_{ij}\right) \tag{2}$$

Where P_i^{DG} and Q_i^{DG}denote DG active and reactive power outputs, while P_i^{BESS} and Q_i^{BESS}represent BESS injections. V_i is the voltage magnitude at bus i, and G_{ij} and B_{ij} denote network conductance and susceptance matrices, respectively.

Network losses are explicitly considered because inappropriate power-flow distribution can reduce the operational benefit obtained from distributed renewable generation. The total feeder loss is determined from the branch resistance and the corresponding active and reactive power flows.

$$P_{load} = \sum_{k=1}^{N_b} R_k \frac{P_k^2 + Q_k^2}{V_k^2} \tag{3}$$

Where R_k represents feeder resistance, while P_k^2 and Q_k^2 denote branch active and reactive power flows.

Under variable renewable output and changing feeder conditions, network losses and voltage deviations can vary considerably. The formulation therefore includes operating constraints to ensure that the resulting control decisions remain compatible with the electrical characteristics of the AND (Huang et al., 2022), (Zhou et al., 2021).

Operating Constraints

The voltage magnitude at every bus is constrained within the prescribed operating interval

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

$$V_i^{min} \leq V_i \leq V_i^{max} \tag{4}$$

The thermal capability of each feeder is also considered

$$S_{ij} \leq S_{ij}^{max} \tag{5}$$

Where S_{ij} is feeder apparent power and S_{ij}^{max} denotes maximum thermal capacity.

Because BESS units are used as controllable resources during both normal operation and restoration, their energy limitations are incorporated into the model. The permissible SOC range is defined by:

$$SOC^{min} \leq SOC_t \leq SOC^{max} \tag{6}$$

The evolution of SOC over consecutive operating periods is determined by the charging and discharging decisions:

$$SOC_{t+1} = SOC_t + \eta_c P_t^{ch} - -\frac{P_t^{dis}}{\eta_d} \Delta t \tag{7}$$

Where $\eta_c P_t^{ch}$ and η_d_dηd denote charging and discharging efficiencies, respectively.

The charging and discharging efficiencies are represented by the corresponding efficiency parameters in the original formulation.

Resilience-Oriented Objective Function

The control problem is formulated to consider both routine operating efficiency and post-contingency recovery. Four performance indicators are incorporated: network power loss, voltage deviation, energy not supplied, and the number of feeder switching operations required during restoration. The resulting multi-objective formulation is expressed as:

$$F = \alpha P_{loss} + \beta V_{loss} + \gamma ENS + \delta SW_{pen} \tag{8}$$

Where *α,β,γ,δ* are weighting coefficients associated with operational objectives.

The voltage deviation index is calculated using:

$$V_{dev} = \sum_{k=1}^{N} \left| V_i - V_{ref} \right| \tag{9}$$

The energy not supplied (ENS) index is subsequently calculated as:

$$ENS = \sum_{t=1}^{T} (P_{demand} - P_{served})\, \Delta t \tag{10}$$

Tthe resilience index is used to quantify the ability of the network to recover its service following a contingency:

$$RI(t) = \frac{P_{served}(t)}{P_{total}} \tag{11}$$

A higher resilience-index value indicates a stronger recovery capability under the considered contingency conditions (Glavic et al., 2021).

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

Topology-Aware Physics-Informed Reinforcement Learning Framework

Framework Concept

The proposed PIRL framework combines three types of information within the control process: electrical operating variables, network topology, and physical operating requirements. This formulation is motivated by the fact that the condition of one bus is not independent of the buses and feeder branches electrically connected to it.

Accordingly, the ADN is first transformed into a graph representation. Topology-related information is then extracted using graph convolution and supplied to the reinforcement learning controller. Physical network requirements are subsequently used when evaluating the consequences of the selected control action. The resulting learning process therefore considers not only whether an action improves the reward but also how that action affects the electrical operating condition of the network. This differs from conventional DRL formulations that generally process network variables as a conventional feature vector without explicitly exploiting feeder connectivity (B. Zhang et al., 2024), (Z. Zhang et al., 2021).

Graph-Based Network Representation

The ADN is naturally described as a graph consisting of buses and feeder connections. In the proposed framework:

$$G = (V, E) \qquad (12)$$

Her,e V represents bus nodes and E denotes feeder connections. Each graph node contains electrical features including voltage magnitude, power flow conditions, DG output, feeder loading, and BESS state of charge.

The connectivity between buses is represented using the adjacency matrix:

$$A_{ij} = \begin{cases} 1, & \text{if buses } i \text{ and } j \text{ are connected} \\ 0, & \text{otherwise} \end{cases} \qquad (14)$$

The graph representation is not limited to binary connectivity. The learning process is interpreted together with branch-related electrical characteristics such as feeder impedance, loading, thermal capability, and switching status. Consequently, branches with identical topological connectivity can still be distinguished according to their electrical characteristics.

An important feature of the proposed formulation is that the topology representation changes whenever feeder switching modifies the network configuration. Thus, the graph presented to the learning agent corresponds to the current network state rather than remaining fixed throughout the entire operation. This property is especially relevant during restoration because fault isolation and reconfiguration can alter the available electrical paths for active- and reactive-power transfer.

Topology-Aware Feature Extraction

Following graph construction, the node information is processed through graph convolution to obtain a representation that reflects both local electrical variables and neighboring network conditions.

$$H^{(l+1)} = \sigma\left(\widehat{D}^{-1/2}\hat{A}\widehat{D}^{-1/2}\right)H^{(l)}W^{(l)} \qquad (15)$$

Where $H^{(l)}$denotes graph feature embeddings, $\hat{A}$ is the augmented adjacency matrix, $\widehat{D}$ represents the degree matrix, and $W^{(l)}$ denotes trainable graph-learning parameters.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

The graph convolution operation aggregates information from electrically connected buses while preserving the structural characteristics of the feeder. Consequently, the representation associated with a given bus is influenced by its own electrical state as well as information propagated from neighboring buses.

This feature becomes particularly useful when the network configuration changes after a fault. Under normal conditions, the graph represents the existing feeder arrangement. After fault isolation and switching, the connectivity matrix changes and the graph-learning process consequently generates a new topology-dependent representation. The reinforcement learning controller can therefore base subsequent decisions on the modified network configuration rather than on the pre-fault structure.

Reinforcement Learning State and Action

The topology-dependent representation is supplied to the reinforcement learning agent, which continuously interacts with the ADN environment. At each decision interval, the agent receives the current electrical and operational information.

The system state is defined by:

$$s_t = [V_t, P_t, Q_t, SOC_t, T_t] \tag{16}$$

Where V_t denotes bus voltage magnitudes, $P_t$and Q_t represent power flow conditions, SOC_t is the battery state of charge, and $T_t$denotes feeder topology status.

The state combines bus voltage information, power-flow conditions, BESS SOC, and feeder topology status. This enables the agent to distinguish operating conditions according to both electrical variables and network configuration.

The control action is expressed as:

$$a_t = [Q_{DG}, P_{BEES}, SW] \tag{17}$$

The action variables provide three complementary control mechanisms. DG reactive-power adjustment contributes to voltage regulation, BESS charging and discharging provide active-power flexibility, and feeder switching changes the network configuration during restoration. Their coordinated use allows the same controller to address both normal operation and contingency recovery.

Policy Optimization

The reinforcement learning policy is trained to maximize the accumulated reward over the operating trajectory:

$$J(\theta) = \mathbb{E}\left[\sum_{t=1}^{T} \gamma^t r_t\right] \tag{18}$$

Where γ denotes the discount factor.

Using cumulative reward allows the controller to account for the future consequences of current decisions. This is particularly important for BESS operation, because charging or discharging at one time step changes the available energy in subsequent periods. A similar consideration applies to feeder restoration, where an immediate switching action can influence the feasibility of later restoration decisions.

The PIRL controller employs a policy-gradient learning strategy. The topology-aware features obtained from the graph convolution layers are combined with the current electrical state, BESS information, and feeder topology before being processed by the policy network. The resulting policy produces control decisions for DG reactive-power regulation, BESS operation, and feeder switching.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

The learning cycle follows a sequential process. The agent first observes the ADN state, generates an action, applies the corresponding control decision, evaluates the resulting network condition, and receives the associated physics-informed reward. The policy parameters are then updated based on the obtained learning signal.

The network architecture contains two graph convolution layers followed by a fully connected policy network. The graph layers capture information exchange between electrically connected buses, while the policy network transforms the resulting representation into the required control decisions.

Physics-Informed Reward Design

A central element of the proposed framework is the incorporation of physical operating information into the learning signal. Without such guidance, the agent may explore actions that produce high short-term rewards but are undesirable from an electrical-system perspective. Examples include excessive network losses, voltage-limit violations, unnecessary switching, or increased ENS during restoration.

To reduce this problem, the PIRL reward considers network loss, voltage deviation, ENS, and switching requirements simultaneously:

$$r_t = -(\alpha P_{loss} + \beta V_{loss} + \gamma ENS + \delta SW_{pen}) \tag{19}$$

The reward formulation therefore links the learning objective directly to the operational performance of the distribution network. Actions producing acceptable voltage conditions, lower losses, reduced unsupplied energy, and fewer unnecessary switching operations receive more favorable learning feedback.

This mechanism also reduces the exploration of undesirable operating regions. Rather than allowing the learning agent to evaluate actions independently of the physical behavior of the ADN, the reward provides explicit information about the electrical consequences of the selected control strategy (Li & Xu, 2024), (Yin et al., 2025).

Physics-Constrained Action Evaluation

Each candidate action generated by the PIRL controller is evaluated using the current operating condition of the distribution network. The action specifies the DG reactive-power output, BESS charging or discharging decision, and feeder switching status. The resulting network state is then evaluated through power-flow analysis.

The feasibility of the action is determined according to the voltage, feeder thermal, power-balance, and BESS SOC constraints defined in (4)-(7). If the selected action produces an undesirable electrical condition, the corresponding penalty is incorporated into the physics-informed reward.

The action evaluation process can be represented as:

$a_t \rightarrow$ *power-flow evaluation* $\rightarrow$ *constraint checking* $\rightarrow r_t$

where a_t denotes the control action generated by the agent and r_t represents the resulting physics-informed reward.

This process establishes a direct connection between the learned policy and the electrical behavior of the ADN. During a feeder contingency, the same evaluation procedure is applied using the post-fault topology. Consequently, switching and DG-BESS decisions are assessed according to the network configuration that actually exists at the corresponding decision step.

Overall PIRL Procedure

The complete PIRL procedure can therefore be viewed as a sequence of topology identification, feature

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

extraction, state observation, control generation, physical evaluation, and policy updating.

The graph component captures the structural characteristics of the ADN. The reinforcement learning component converts the observed network condition into coordinated DG-BESS and switching decisions. The physics-informed mechanism evaluates the electrical consequences of these decisions and supplies the corresponding learning feedback.

Because the graph is updated when the feeder configuration changes, the controller can adapt its decision process to renewable variability as well as to post-fault topology changes. The overall sequence is presented in Figure 1.

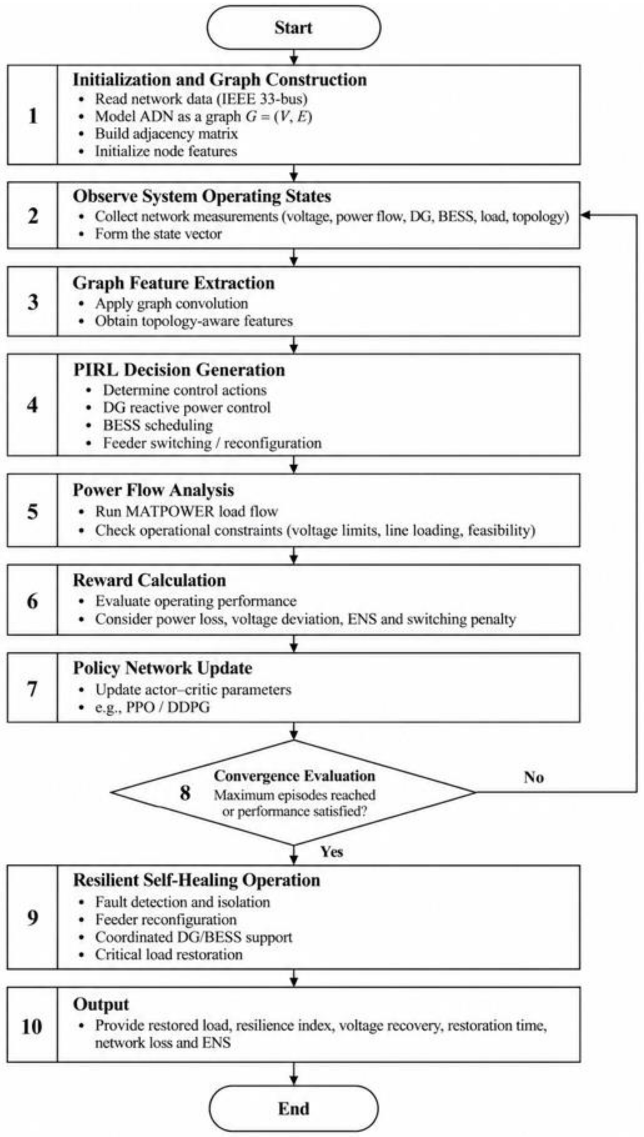

Figure 1. Flowchart of the proposed PIRL

RESULTS AND DISCUSSION

Simulation Setup

The proposed framework is assessed using the IEEE 33-bus active distribution network (ADN). Three PV-based distributed generation units are integrated at buses 6, 18, and 30, while BESS units are installed at buses 14 and 25. In addition, controllable tie-switches are incorporated into the network to support feeder reconfiguration when abnormal operating conditions occur. The system has a total demand of 3.72 MW and

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

2.30 MVAr, and the installed PV generation corresponds to a penetration level of 42% relative to the total load. Each BESS is rated at 500 kWh with a charging and discharging efficiency of 95%.

For the learning process, the PIRL agent is trained for 1000 episodes using an adaptive policy optimization strategy. The learning structure comprises two graph convolutional layers connected to a fully connected policy network. Through topology-aware feature aggregation, the model captures the relationships between electrically interconnected buses and incorporates these relationships into the decision-making process. The main simulation and learning settings are listed in Table 1.

Table 1. Simulation Parameters of the Proposed PIRL Framework

Parameter	Value
Test system	IEEE 33-bus ADN
Total load demand	3.72 MW, 2.30 MVAr
PV penetration	42%
Number of PV units	3
PV locations	Buses 6, 18, and 30
Number of BESS units	2
BESS locations	Buses 14 and 25
BESS capacity	500 kWh
Voltage limits	0.95 - 1.05 p.u.
BESS SOC limits	20 - 90%
BESS efficiency	95%
Training episodes	1000
Learning rate	0.0005
Discount factor	0.98
Graph convolution layers	2
Policy network	Fully connected

The selected configuration enables the proposed approach to be examined under operating conditions involving distributed renewable generation, energy storage, and network topology changes. The PV units introduce variability in renewable power availability, whereas the BESS units provide controllable energy resources for balancing and restoration support. Meanwhile, the controllable tie-switches allow the learning agent to modify feeder connectivity when required, thereby providing an operational mechanism for adaptive network restoration and resilient operation.

Because reinforcement learning is sensitive to random initialization and exploration, each learning configuration is trained using multiple independent random seeds. The reported performance metrics are obtained from the independent runs and are presented as mean ± standard deviation. The same number of independent runs and the same training budget are used for all learning-based benchmark methods to ensure a consistent comparison.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

Contingency and Uncertainty Scenarios

Four representative scenarios are considered by combining PV variability, load changes, and feeder contingencies. Following each fault, the affected section is isolated and feeder reconfiguration is performed using available tie-switches, with DG and BESS providing additional support. The same scenarios and training conditions are applied to all methods.

Table 2. Contingency and Uncertainty Scenarios

Scenario	Faulted Branch	PV Condition	Load Condition
S1	9–10	100%	100%
S2	14–15	60%	120%
S3	19–20	90%	100%
S4	25–26	40–90%	130%

Voltage Regulation Performance

Voltage regulation is evaluated under the four operating scenarios in Table 2, which combine feeder faults with different PV penetration levels and load conditions. Figure 2 presents the voltage profiles of the IEEE 33-bus system for the four control methods. The voltage limits are maintained at 0.95–1.05 p.u.

Table 3. Voltage Regulation Performance

Method	Minimum Voltage (p.u.)	Maximum Voltage Deviation (p.u.)	Within Limits (%)
DRL-based	0.951	0.049	89.6
Graph-Based Volt-VAR	0.958	0.042	94.8
PI-MADRL	0.968	0.032	97.5
PIRL	0.976	0.024	99.1

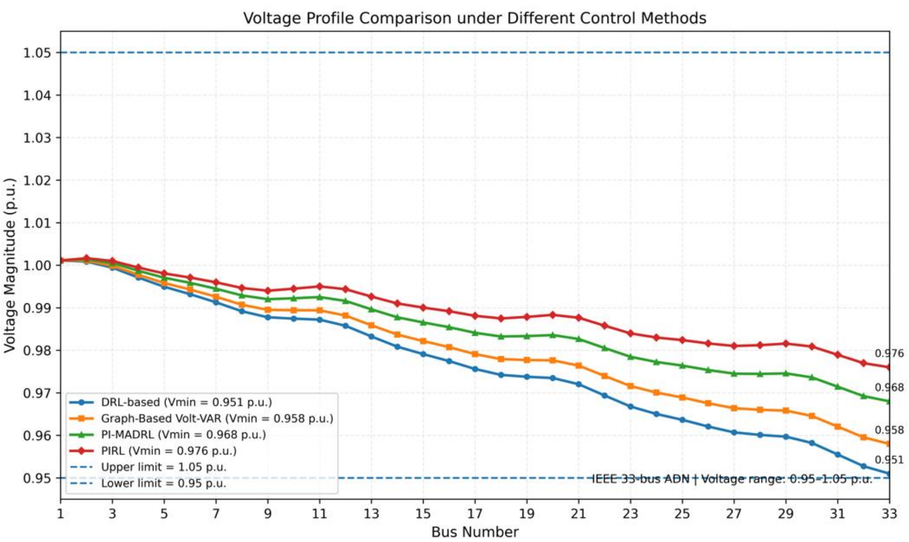

Figure 2. Voltage profile comparison under different control methods.

As shown in Figure 2, the DRL-based method exhibits the largest voltage drop toward the downstream buses, with a minimum voltage of 0.951 p.u. Graph-Based Volt-VAR improves the profile to 0.958 p.u.,

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

while PI-MADRL achieves 0.968 p.u. The proposed PIRL provides the highest minimum voltage of 0.976 p.u. and maintains a smoother voltage profile along the feeder.

These results are consistent with Table 3, where the maximum voltage deviation decreases from 0.049 p.u. for DRL-based control to 0.042, 0.032, and 0.024 p.u. for Graph-Based Volt-VAR, PI-MADRL, and PIRL, respectively. The percentage of operating points within the allowable voltage range increases from 89.6% to 94.8%, 97.5%, and 99.1%.

The improved performance of PIRL is attributed to the combined consideration of network topology, DG-BESS coordination, and feeder reconfiguration. This enables the controller to respond more effectively to PV fluctuations, increased loading, and post-fault topology changes, thereby improving voltage security during both normal and contingency conditions.

Network Power Loss Reduction

Total network power loss is used to evaluate the operating efficiency of the IEEE 33-bus ADN under the considered operating conditions. The comparison is performed using the same PV, load, and contingency scenarios for all control methods.

Table 4. Network Power Loss Comparison

Method	Power Loss (kW)	Loss Reduction (%)
DRL-based	160.4	26.0
Robust DG-BESS	145.8	32.7
Graph-Based Volt-VAR	131.6	39.3
PIRL	121.3	44.0

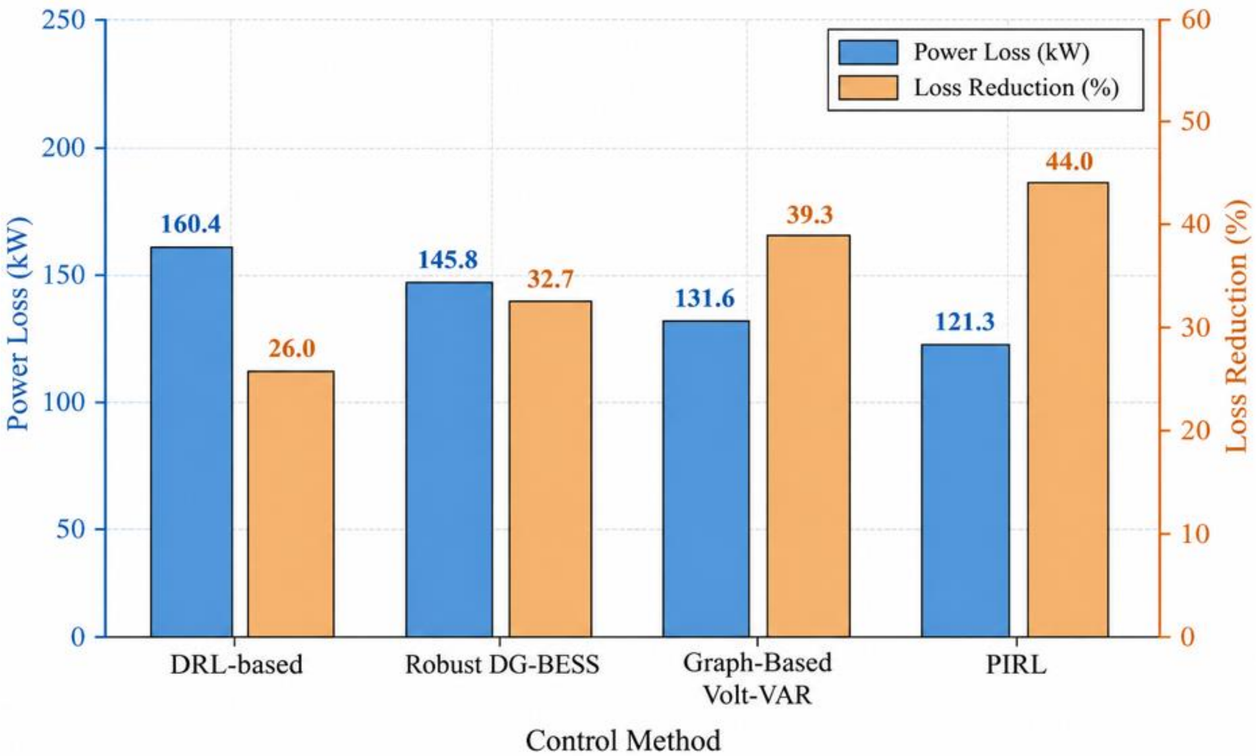

Figure 3. Network power loss comparison under different control methods.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

As shown in Table 4 and Figure 3, the DRL-based method results in a network loss of 160.4 kW, corresponding to a 26.0% reduction from the base case. Coordinating DG and BESS reduces the loss further to 145.8 kW, while the Graph-Based Volt-VAR method achieves 131.6 kW.

The proposed PIRL framework obtains the lowest loss of 121.3 kW, representing a 44.0% reduction. This improvement is mainly associated with topology-aware decision making and coordinated DG-BESS operation, which enable more effective power-flow redistribution under varying generation, loading, and network configurations. The results also indicate that improved voltage regulation and reduced feeder loading contribute jointly to the reduction in network losses.

Resilient Self-Healing Performance

The restoration time T_rec is measured from the fault occurrence instant to the first time instant at which the interrupted loads are fully restored and all operating constraints are simultaneously satisfied. In the present study, the feeder fault is introduced at t = 4 h; therefore, the restoration time reported in Table 5 represents the elapsed post-fault time rather than the absolute simulation time.

Table 5 Figure 4 presents the load restoration trajectories following the feeder fault. The Hybrid DL, DRL-based, and PI-MADRL methods restore the interrupted load within 30, 23, and 18 min, respectively. The proposed PIRL framework completes the restoration in 13 min, indicating a faster response to the post-fault network condition.

Table 5. Self-Healing Restoration Performance

Method	Restoration Time (min)	ENS (kWh)	Resilience Index
Hybrid DL	30	205	0.84
DRL-based	23	150	0.89
PI-MADRL	18	105	0.93
PIRL	13	72	0.96

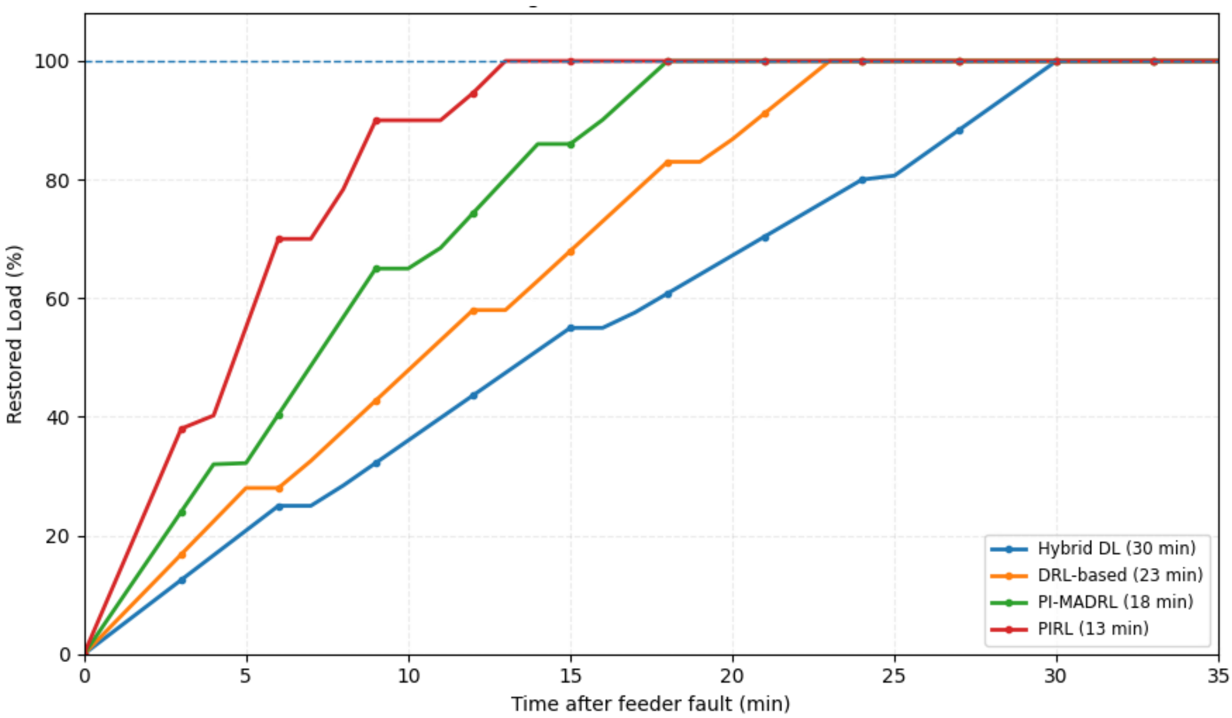

Figure 4. Self-healing load restoration performance under the feeder fault.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

The reduction in restoration time is accompanied by a decrease in ENS from 205 kWh for Hybrid DL to 72 kWh for PIRL. At the same time, the resilience index increases from 0.84 to 0.96, reflecting improved recovery performance after the feeder outage.

The improved response of PIRL is mainly related to its topology-aware decision process and coordinated DG-BESS operation. Following the fault, the controller can adapt the network configuration and utilize distributed resources to support load restoration. This enables power to be redistributed through feasible feeder paths while reducing the duration of supply interruption.

Overall, the results demonstrate that PIRL can provide faster and more effective load restoration under the considered feeder contingency. The combination of adaptive topology handling and coordinated DG-BESS support contributes to improved resilience and reduced energy interruption.

Learning Convergence Analysis

The convergence behavior of the three learning-based methods is evaluated over 1000 training episodes under the same operating scenarios. Figure 5 shows the evolution of normalized cumulative reward during training.

Table 6. Learning Convergence Comparison

Method	Stable Convergence Episode	Final Normalized Reward
DRL-based	760	0.78
PI-MADRL	610	0.86
PIRL	560	0.91

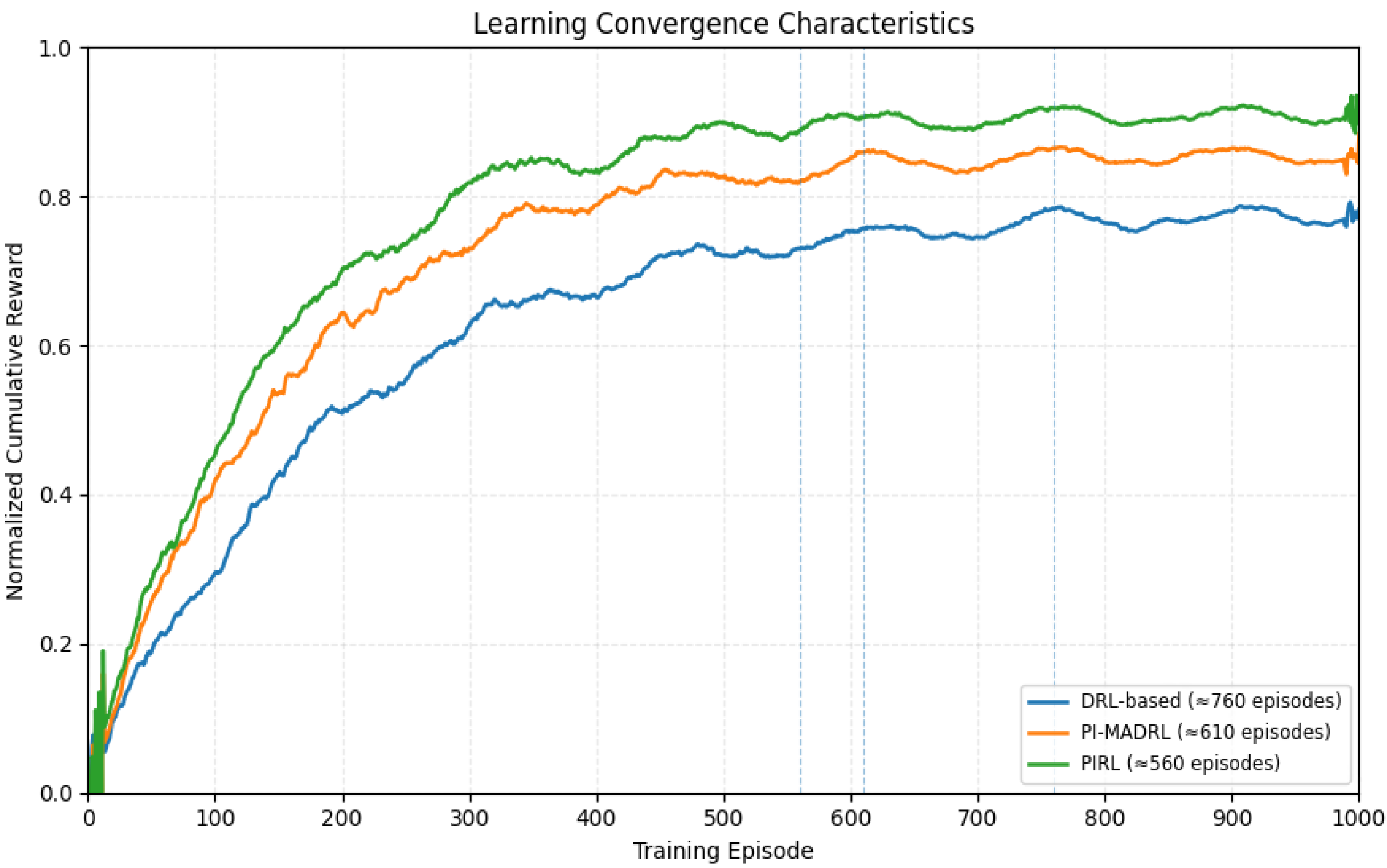

Figure 5. Learning convergence characteristics of different control methods.

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

As shown in Figure 5, the DRL-based method requires approximately 760 episodes to reach a stable policy, while PI-MADRL and PIRL converge at about 610 and 560 episodes, respectively. PIRL also achieves the highest final normalized reward of 0.91.

The faster convergence of PIRL is mainly attributed to the combination of topology-aware state representation and physics-informed reward guidance. The former captures network connectivity, while the latter penalizes undesirable voltage, loss, ENS, and switching conditions during exploration. Consequently, PIRL can reduce ineffective exploration and learn a stable control policy with fewer training episodes.

For comparison across learning configurations, the cumulative episode reward is normalized using the predefined reward range obtained from the training environment. The normalized reward is reported on a dimensionless scale, where higher values indicate better overall operating performance. The same normalization procedure is applied to all benchmark methods.

Ablation Analysis of PIRL Components

To further examine the contribution of the main mechanisms incorporated into the proposed PIRL framework, a component-based analysis is conducted using the control methods already considered in this study. The comparison focuses on the progressive inclusion of learning capability, topology-aware information, physics-informed decision guidance, and coordinated DG-BESS operation. The DRL-based method is used as the conventional learning-based reference, while Graph-Based Volt-VAR represents topology-aware control. PI-MADRL provides a physics-informed multi-agent learning reference, and Robust DG-BESS represents coordinated distributed-resource operation. The complete PIRL framework integrates these control characteristics together with adaptive feeder reconfiguration.

All methods are evaluated using the same IEEE 33-bus distribution network, operating constraints, PV and load conditions, and contingency scenarios described previously. The same performance indicators are considered, including restoration time, energy not supplied (ENS), network power loss, and resilience index. This common evaluation procedure allows the effect of the additional mechanisms incorporated in PIRL to be examined without changing the underlying test environment.

The analysis is summarized in Table 7. The DRL-based method provides the reference learning-based performance without the explicit graph-based and physics-informed mechanisms used by PIRL. Graph-Based Volt-VAR introduces network-structure information into the control process, whereas PI-MADRL represents the physics-informed learning approach considered in the study. Robust DG-BESS explicitly emphasizes coordinated operation of distributed generation and battery storage. PIRL combines these aspects within a unified topology-aware and physics-informed reinforcement-learning framework.

Table 7. Component-Based Analysis of the Considered Control Methods

Method	Learning-Based Control	Topology Awareness	Physics-Informed Guidance	DG-BESS Coordination	Restoration Time (min)	ENS (kWh)	Power Loss (kW)	Resilience Index
Hybrid DL	✓	–	–	–	30	205	—	0.84
DRL-based	✓	–	–	–	23	150	160.4	0.89
Robust DG-BESS	–	–	–	✓	—	—	145.8	—
Graph-Based Volt-VAR	✓	✓	–	—	—	—	131.6	—

ISSN 2278-2540 | DOI: 10.51583EMAS | Volume XV, Issue IX September 2026

PI-MADRL	✓	—	✓	—	18	105	—	0.93
PIRL	✓	✓	✓	✓	13	72	121.3	0.96

Discussion

The results in Table 7 indicate that the performance improvement of PIRL is associated with the combined use of the mechanisms incorporated in the proposed framework. The DRL-based method requires 23 min for restoration, with an ENS of 150 kWh and a resilience index of 0.89. Incorporating topology-aware control through the Graph-Based Volt-VAR approach reduces the network power loss to 131.6 kW, compared with 160.4 kW for the DRL-based method. The PI-MADRL method further demonstrates the benefit of physics-informed learning, achieving a restoration time of 18 min, an ENS of 105 kWh, and a resilience index of 0.93.

The Robust DG-BESS result also confirms the importance of coordinated distributed-resource operation, reducing the network power loss to 145.8 kW. In the proposed PIRL framework, topology-aware learning, physics-informed guidance, coordinated DG-BESS operation, and feeder reconfiguration are integrated within the same decision-making process. Consequently, PIRL achieves 13 min restoration time, 72 kWh ENS, 121.3 kW network power loss, and a resilience index of 0.96.

These comparisons suggest that the performance of PIRL does not originate from reinforcement learning alone. Rather, the results support the importance of incorporating network topology, physical operating information, and coordinated DG-BESS control into the learning-based restoration process. The proposed framework therefore provides a unified approach for adapting the control policy to changing feeder configurations and renewable operating conditions.

Training and Inference Efficiency

Beyond control quality, computational efficiency is important for practical deployment in active distribution networks. The learning models are trained offline, whereas the trained policy is executed online to determine control actions from the current network state. Therefore, training and inference times are evaluated separately.

As summarized in Table 8, the DRL-based method requires 56.4 ± 1.8 min for training and converges after approximately 780 episodes. PI-MADRL reduces the training time to 49.8 ± 1.5 min and converges at 630 episodes. PIRL requires 45.6 ± 1.3 min and reaches stable convergence after approximately 560 episodes. Although PIRL has a slightly higher inference time of 12.4 ± 0.6 ms due to its topology-aware processing, this computational requirement remains suitable for online control.

Table 8. Computational Performance

Method	Training Time (min)	Stable Convergence (episodes)	Inference Time (ms)
DRL-based	56.4 ± 1.8	780	7.8 ± 0.4
PI-MADRL	49.8 ± 1.5	610	10.6 ± 0.5
PIRL	45.6 ± 1.3	560	12.4 ± 0.6

The results indicate a trade-off between training efficiency and online inference cost. PIRL requires slightly more computation during inference because of its structured network representation, but it achieves faster policy convergence and lower overall training time. The inference time of 12.4 ms is sufficiently small compared with typical distribution-network control intervals, supporting the feasibility of PIRL for near-real-time adaptive operation.

DISCUSSION

The results indicate that PIRL improves the overall operational and resilience performance of the renewable-integrated ADN rather than optimizing a single operating criterion. The proposed framework achieves improved voltage security, lower network losses, and faster post-fault restoration, while requiring fewer training episodes to obtain a stable policy.

The main advantage stems from the integration of topology-aware representation and physics-informed decision guidance. The former enables the agent to capture changes in feeder connectivity, particularly after fault isolation and reconfiguration, whereas the latter directs the learning process toward electrically feasible operating states. Coordinated DG-BESS control then provides the flexibility required to regulate voltage, redistribute power flows, and support load restoration.

Importantly, the consistent improvement across voltage regulation, loss reduction, self-healing, and convergence suggests that these components act synergistically rather than independently. Thus, PIRL provides a unified control framework capable of adapting to renewable variability and changing network configurations while maintaining the essential operating requirements of the ADN.

Scope and Practical Interpretation

The results presented in this study constitute a simulation-based evaluation of the proposed learning framework rather than a field demonstration of autonomous distribution-system self-healing. The IEEE 33-bus network, renewable profiles, load variations, feeder contingencies, and switching operations are represented in a simulation environment, and the reported restoration performance therefore demonstrates the capability of the proposed controller under the considered modeled conditions.

In particular, the reported restoration time represents the control and network-restoration response within the simulation environment and does not include physical protection-device operation, communication delays, field switching delays, cyber-security constraints, or hardware-in-the-loop effects. Therefore, the results should be interpreted as evidence of the feasibility and potential of the proposed PIRL framework for resilient ADN operation rather than as direct validation of autonomous field deployment.

Further validation using larger distribution networks, real feeder measurements, hardware-in-the-loop experiments, and realistic protection and communication models would be required before practical deployment.

CONCLUSION

The proposed PIRL framework provides a simulation-based approach for resilient operation and adaptive restoration of renewable-integrated active distribution networks. By combining topology-aware graph representation, physics-informed reward guidance, coordinated DG-BESS control, and feeder reconfiguration, the framework links reinforcement-learning decisions with the electrical and structural characteristics of the distribution network. The IEEE 33-bus simulation results indicate that the proposed approach can maintain improved voltage performance, reduce network losses, and accelerate load restoration under the considered combinations of PV variability, load changes, and feeder contingencies. The restoration analysis further shows that the coordinated use of network reconfiguration and distributed energy resources can reduce ENS and improve the calculated resilience index. Nevertheless, these results represent a simulation-based proof of concept rather than a field demonstration of autonomous self-healing. The present study is limited to the IEEE 33-bus system and does not explicitly model physical protection coordination, communication delays, hardware switching constraints, or hardware-in-the-loop operation. Future work will therefore focus on larger-scale distribution systems, broader uncertainty conditions, independent multi-run statistical evaluation, and experimental or hardware-in-the-loop validation. Overall, the study demonstrates the potential of topology-aware and physics-informed reinforcement learning as a unified decision-making framework for resilient operation of renewable-rich active distribution networks.

REFERENCES

1. Behzadi, S., Bagheri, A., & Rabiee, A. (2024). Optimal operation of reconfigurable active distribution networks aiming at resiliency improvement. Electric Power Systems Research, 230, 110169. https://doi.org/10.1016/j.epsr.2024.110169.
2. Byeon, G., & Kim, K. (2024). Distributionally robust decentralized Volt-Var control with network reconfiguration. IEEE Transactions on Smart Grid, 15(5), 4705–4718,. https://doi.org/10.1109/TSG.2024.3377910.
3. Dominguez-Garcia, A. D., Zholbaryssov, M., Amuda, T., & Ajala, O. (2024). An online feedback optimization approach to voltage regulation in inverter-based power distribution networks. IEEE Transactions on Power Systems, 39(1), 1145–1156, https://doi.org/10.1109/TPWRS.2023.3304112.
4. Glavic, M. (2019). Deep) reinforcement learning for electric power system control and related problems: A short review and perspectives. Annual Reviews in Control, 48, 22–35, https://doi.org/10.1016/j.arcontrol.2019.09.008.
5. Glavic, M., Moreno, R., Bi, T., & Shahidehpour, M. (2021). Reinforcement learning for electric power system decision and control: Past considerations and perspectives. IEEE Transactions on Smart Grid, 12(6), 5128–5142, https://doi.org/10.1109/TSG.2021.3070934.
6. Huang, Y., Yang, Q., Tan, J., & Ukil, A. (2022). Deep reinforcement learning for real-time energy management in smart grids. Energy Reports, 8, 1245–1257, https://doi.org/10.1016/j.egyr.2022.08.190.
7. Li, B., & Xu, Q. A. (2024). A machine learning-assisted distributed optimization method for inverter-based Volt-VAR control in active distribution networks. IEEE Transactions on Power Systems, 39(2), 2668–2681, https://doi.org/10.1109/TPWRS.2023.3279303.
8. Qiu, W., Yadav, A., You, S., Dong, J., Kuruganti, T., Liu, Y., & Yin, H. (2024). Neural networks-based inverter control: Modeling and adaptive optimization for smart distribution networks. IEEE Transactions on Sustainable Energy, 15(2), 1039–1049, https://doi.org/10.1109/TSTE.2023.3324219.
9. Wang, S., Hou, Y., Guan, X., Liu, S., & Huo, Z. (2025). Resiliency-informed optimal scheduling of smart distribution network with urban distributed photovoltaic: A stochastic P-robust optimization. Energy, 313, 133449. https://doi.org/10.1016/j.energy.2024.133449.
10. Xu, Y., Dong, Z. Y., Zhang, R., & Hill, D. J. (2017). Multi-timescale coordinated voltage/VAR control of high renewable-penetrated distribution systems. IEEE Transactions on Power Systems, 32(6), 4398–4408, https://doi.org/10.1109/TPWRS.2017.2659779.
11. Yin, C., Dong, J., & Zhang, Y. (2025). Distributionally robust bilevel optimization model for distribution network with demand response under uncertain renewables using Wasserstein metrics. IEEE Transactions on Sustainable Energy, 16(1), 118–130, https://doi.org/10.1109/TSTE.2024.3509314.
12. Zhang, B., Cao, D., Hu, W., Ghias, A. M. Y. M., & Chen, Z. (2024). Physics-informed multi-agent deep reinforcement learning enabled distributed voltage control for active distribution network using PV inverters. International Journal of Electrical Power & Energy Systems, 155, 109641. https://doi.org/10.1016/j.ijepes.2023.109641.
13. Zhang, C., Xu, Y., Dong, Z. Y., & Wong, K. P. (2018). Robust coordination of distributed generation and energy storage in distribution networks. IEEE Transactions on Smart Grid, 9(2), 1488–1499, https://doi.org/10.1109/TSG.2016.2583461.
14. Zhang, Z., Wang, J., & Chen, B. (2021). Deep reinforcement learning for energy management in microgrids. Applied Energy, 288, 116618. https://doi.org/10.1016/j.apenergy.2021.116618.
15. Zhao, J., Zheng, T., & Litvinov, E. (2016). A unified framework for defining and measuring flexibility in power system. IEEE Transactions on Power Systems, 31(1), 339–347, https://doi.org/10.1109/TPWRS.2015.2390039.
16. Zhou, K., Yang, S., & Shao, Z. (2021). Digital twin-based optimization for energy systems. Applied Energy, 285, 116419. https://doi.org/10.1016/j.apenergy.2020.116419.